

Virtuelle Bibliotheken flexibel gestalten

Wiebke Oeltjen

This document appeared in

Bernhard Mittermaier (Eds.):

eLibrary - den Wandel gestalten

5. Konferenz der Zentralbibliothek

Proceedings of the WissKom 2010: 5. Konferenz der Zentralbibliothek, 08.-11. November 2010, Jülich

Schriften des Forschungszentrums Jülich / Reihe Bibliothek/Library, Vol. 20

Zentralbibliothek (ZB)

Forschungszentrum Jülich GmbH, Zentralbibliothek, Verlag, 2010

ISBN: 978-3-89336-668-2

Virtuelle Bibliotheken flexibel gestalten

Wiebke Oeltjen

Zusammenfassung

In dem Beitrag werden virtuelle Bibliotheken vorgestellt, die auf der Open-Source-Software MyCoRe basieren. Das Spektrum an Einsatzbereichen wird an verschiedenen Beispielen gezeigt. Dazu gehören sowohl Spezialbestände, wie z.B. orientalische Handschriften und digital erfasste und katalogisierte Abbildungen von Papyri als auch Zeitschriften und klassische Bibliotheksbestände. Weiter wird die Funktionalität und Flexibilität von MyCoRe beschrieben.

Abstract

This contribution introduces virtual libraries, based on the open source software MyCoRe. By means of several examples the wide range of applications is shown, including special content of the virtual libraries for instance oriental manuscripts and digitized and electronically catalogued papyrus collections as well as journals and conventional library content. Furthermore the functionality and flexibility of MyCoRe is described.

Einleitung

Wissenschaftliche Forschung braucht uneingeschränkten Zugang zu Forschungsmaterial und wissenschaftlichen Publikationen. Virtuelle Bibliotheken eröffnen solch einen Zugang zu kostbaren Handschriften, Papyri und anderen historischen Dokumenten ebenso wie zu klassischen Bibliotheksbeständen und Katalogen. Der virtuelle Zugang schützt einerseits die Objekte, das physische Material und ermöglicht andererseits weitergehende Forschung. Die Virtualisierung des Forschungsmaterials ist der erste Schritt zur Bereitstellung, die bibliothekarische Erschließung und Erfassung von Metadaten ein weiterer (Oeltjen 2007). Ein Dokumentenserver, eine Archivsoftware ermöglicht das Erfassen, Bereitstellen, Recherchieren und Zugreifen auf die Daten via Internet. Das ist die virtuelle Bibliothek.

Das Open-Source-Softwareprojekt MyCoRe¹ (Lützenkirchen 2002) stellt einen Softwarekern zur Verfügung, der die Basis bildet für eine Reihe von Anwendungen², die

¹ Der Name „MyCoRe“ ist ein Akronym für „My Content Repository“. Die Aussprache „my core“ weist darauf hinweisen, dass es sich um einen Softwarekern handelt.

² Siehe „Anwendungen“ auf der MyCoRe-Homepage unter <http://www.mycore.de>.

als virtuelle Bibliotheken dienen. Dazu gehören digitale Bibliotheken mit großen Datenmengen wie z.B. die „Digitale Bibliothek für Thüringer Hochschulen“³ (DBT) oder der Dokumenten- und Publikationsserver der Universität Duisburg-Essen „DuEPublico“⁴. Außerdem werden verschiedene Spezialbestände mit MyCoRe-Anwendungen virtuell bereitgestellt und zugänglich gemacht. Verwaltet werden z.B. alte Bibliotheksbestände mit orientalischen Manuskripten und Papyri. Desweiteren sind Online-Journals und ein Zeitschriftenportal mit MyCoRe realisiert. Im Folgenden sollen verschiedene solcher MyCoRe-Anwendungen vorgestellt werden. Anschließend wird der Funktionsumfang der Software MyCoRe skizziert. Gezeigt werden soll weiter, dass durch eine modulare Bauweise der Software virtuelle Bibliotheken mit MyCoRe flexibel gestaltet werden können.

Spezialbestände in virtuellen Bibliotheken

Auf der Basis der Software MyCoRe sind bundesweit verschiedene virtuelle Bibliotheken produktiv im Einsatz. Dazu gehören Archive, Bilddatenbanken und Kataloge, die einen Online-Zugang zu Forschungsmaterial, Publikationen und Dokumenten verschiedener Art gewähren. Im Folgenden sollen Beispiele für virtuelle Bibliotheken mit Spezialbeständen genannt werden.

Orientalische Handschriften

An der Universität Leipzig wurden in den letzten Jahren eine Reihe virtueller Bibliotheken für orientalische Handschriften entwickelt, die auf der Software MyCoRe basieren (Hanstein, Karmun, Kupferschmidt 2010), wie z.B. die „Damaszener Familienbibliothek Refaiya“⁵, das „Portal der Handschriften Javas“⁶ oder die „Islamischen Handschriften der Universitätsbibliothek Leipzig“⁷, um nur einige zu nennen. Bei der Implementierung wurden Erfahrungen gesammelt, die für vergleichbare Projekte mit Mehrsprachigkeit und Mehrschriftlichkeit von Bedeutung sein können. So wurden die arabischen, persischen und türkischen Schriften mit dem Unicode Standard UTF-8 in die Datenbank integriert, so dass die Metadaten in der jeweiligen Schrift im Browser dargestellt werden können. Die Beschreibungen der Manuskripte und Handschriften sind in verschiedenen Schriften und Sprachen als Metadaten erfasst.

³ Siehe <http://www.db-thueringen.de>

⁴ Siehe <http://duepublico.uni-duisburg-essen.de>

⁵ Siehe <http://refaiya.dl.uni-leipzig.de>

⁶ Siehe <http://www.manuscripts-java.org>

⁷ Siehe <http://www.islamic-manuscripts.net>

Eine Recherche in den Metadaten kann ebenfalls in der jeweiligen Sprache mit der entsprechenden Schrift vorgenommen werden. Als Eingabehilfe für die orientalischen Schrift- und Sonderzeichen wurde eine Bildschirmtastatur entwickelt und eingebunden (siehe Abb. 1). Sie enthält „neben den erforderlichen Transkriptionssonderzeichen auch Buchstaben und Wörter in Arabisch, Persisch, (Osmanisch-)Türkisch und Jawi⁸“ (Hanstein, Kupferschmidt, Scharsky 2009).



Abb. 1: Bildschirmtastatur, arabisch (Ausschnitt)

Eine besondere Herausforderung stellt die rechts-links-Schreibrichtung (*right-to-left*) und die „Mehrschriftlichkeit“ dar (ebda.). Dies wurde in den digitalen Bibliotheken für orientalische Handschriften so implementiert, dass auch ein Wechsel zwischen den Sprachen (Deutsch, Englisch, Arabisch o.a.) möglich ist. Darüber hinaus werden beschreibende Texte bidirektional angezeigt (siehe Abb. 2).

Titel :	
↳ wie in Hs. :	Bl. 1v: al-Ġuz' al-auwal min Sīrat as-saīyidnā rasūl Allāh
	ق 1: الجرز الأول من سيرة سيدنا رسول الله

Abb. 2: Bidirektionale Schreibrichtung

Neben den beschreibenden Metadaten sind auch Abbildungen von Manuskripten, handgeschriebenen Büchern, Briefen oder anderen Publikationen in diesen virtuellen Bibliotheken enthalten. So ist ein Zugang zu historisch bedeutsamen Bibliotheksbeständen geschaffen worden, der nun der Forschung via Internet zur Verfügung steht.

⁸ Malaiisch mit arabischen Buchstaben geschrieben.

Verteilte Suche auf dem Papyrus-Portal

Das Papyrus-Portal⁹ ist ein virtueller Zusammenschluss von digitalisierten und elektronisch katalogisierten Papyrussammlungen die bundesweit in zehn Bibliotheken vorliegen, nämlich Bonn, Erlangen, Gießen, Halle, Heidelberg, Jena, Köln, Leipzig, Trier und Würzburg. Ziel des Portals ist es, eine zentrale Recherche zu ermöglichen und eine einheitliche Präsentation der Suchergebnisse zu bieten. Da die Datenbestände an den jeweiligen Standorten auf unterschiedlichen Plattformen und Datenbanken vorliegen, wurden verschiedene Schnittstellen für die Recherche-Zugriffe auf die Datenbanken implementieren. MyCoRe selbst ermöglicht eine „Suche via Web-Service in datenmodellgleichen entfernten Anwendungen“, darüber hinaus wurden Recherchen auf FileMaker Datenbanken mittels JDBC-Treiber realisiert, und eine weitere Schnittstelle nutzt den „Zugriff über eine Open-Archive-Schnittstelle auf eine Allegro/Hans-Datenbank“, um relevante Suchinformationen zu extrahieren (Freitag et al. 2009, S. 15f.). Zu erwähnen ist noch, dass der vereinheitlichte Zugriff auf die verteilten Daten und Bibliotheksbestände ein standardisiertes Datenmodell verwenden und einen systematischen Zugriff über Klassifikationen erlauben.

Große Datenmengen im Zeitschriftenportal

Ein weiteres Beispiel für eine virtuelle Bibliothek auf der Basis der Software MyCoRe ist das Zeitschriftenportal „Journals@UrMEL“ an der Thüringer Universitäts- und Landesbibliothek (ThULB¹⁰) in Jena. Auf dem Zeitschriftenserver werden mehr als 450 E-Journals, retrodigitalisierte historische Zeitschriften und Kalenderblätter digital publiziert. Die einzelnen Zeitschriften unterscheiden sich dabei im Erscheinungsbild, weisen aber jeweils das gleiche Datenmodell auf und haben eine eigene Webseitenstruktur und Rechteverwaltung. Technisch sind alle Zeitschriften und Kalenderblätter Bestandteil einer MyCoRe-Anwendung. Es können mit geringem administrativem Aufwand neue Zeitschriften per Mausklick hinzugefügt werden. Die Zeitschriften werden teilweise mit Zugangsbeschränkung veröffentlicht, teils sind die Artikel im freien Zugriff abrufbar (*Open Access*). Der Umfang an gespeicherten Daten mit Dokumenten im PDF-Format zeigt die Leistungsfähigkeit eines MyCoRe-Systems, auch wenn große Datenmengen (im Terabyte-Bereich) vorgehalten und bereitgestellt werden.

⁹ Siehe <http://www.papyrusportal.de>

¹⁰ Siehe <http://zs.thulb.uni-jena.de>

Klassische Bibliotheksbestände

Neben den MyCoRe-Anwendungen mit den verschiedenen Spezialbeständen sind die digitalen Bibliotheken und Dokumenten- und Publikationsserver im universitären Kontext auf der Basis von „Miless“ zu nennen. Miless ist die Ursprungssoftware, aus der heraus MyCoRe entwickelt wurde (Lützenkirchen 2002). Der Schwerpunkt dieses Softwarezweiges liegt im Bereich universitärer Dokumentenserver zur Bereitstellung von multimedialen Lehr- und Lernmaterialien mit Animationen, Simulationen, Audio- und Videomaterial, sowie elektronische Publikationen und Dissertationen. Darüber hinaus können Semesterapparate eingerichtet werden, die digitalisierte Texte, Zeitschriftenaufsätze, Buchauszüge, Literaturempfehlungen, Internet-Links und andere Materialien zu einer Lehrveranstaltung bereitstellen. Die digitale Bibliothek „DuEPublico“ (*Duisburg-Essen Publications online*) ist hier exemplarisch zu nennen. Dieser Dokumentenserver ist zertifiziert nach den Kriterien der „Deutschen Initiative für Netzwerkinformation“ (DINI). Er erfüllt damit die Anforderungen an Dokumenten- und Publikationsservices, wie sie in dem Kriterienkatalog beschrieben wurden (DINI 2007). Der Hochschulschriftenserver bietet die Möglichkeit, Publikationen im Volltext zu archivieren und im Sinne des „Open Access“ über das Internet frei zugänglich zu machen. In der Bundesrepublik wird Miless inzwischen an neun Universitäten und Universitätsbibliotheken eingesetzt.

Die Funktionalität von MyCoRe

Um zu einer eigenen digitalen Bibliothek auf der Basis von MyCoRe zu gelangen, bietet es sich an, ausgehend von der Beispielanwendung DocPortal, die auf MyCoRe basiert, eigene Anpassungen vorzunehmen. MyCoRe bietet eine große Flexibilität in der Anpassbarkeit. Dazu gehört die Möglichkeit, das Datenmodell zur Beschreibung der Objekte mit Metadaten so an die eigenen Anforderungen anzupassen, dass individuelle Datenbestände mit beschreibenden Metadaten erfasst werden können. Ein MyCoRe-Metadaten-Objekt, das intern im XML-Format vorliegt, besteht aus drei Teilen. Das sind zuerst strukturelle Informationen, die zur Einordnung des Datensatzes in eine logische Struktur dienen, dann die eigentlichen Metadaten auf der Grundlage des Datenmodells und drittens die Service-Daten, die Informationen zur Verwaltung des Metadaten-Satzes wie Erstellungsdatum oder Zugriffsrechte enthalten.

Die Erfassung der Metadaten wird über Eingabeformulare realisiert, die an das eigene Datenmodell anzupassen sind. Eine Benutzerverwaltung regelt den Zugang zum System und erlaubt die gezielte Freischaltung von Bearbeitungsformularen und Arbeitsabläufen. Ebenso können Zugriffsrechte für das Lesen, Bearbeiten und Löschen von Dokumenten und Dateien einzelnen Benutzern oder bestimmten Gruppen zugeordnet werden. Die Zuordnung von Rechten zu Objekten erfolgt über Access Control Listen (ACL). Alternativ kann auch auf der Basis von IP-Adressbereichen Zugang zu Dokumenten oder anderen Objekten gewährt werden.

Für die inhaltliche Erschließung von Bibliotheksbeständen ist es unabdingbar, dass die Objekte klassifiziert werden. Die verschiedenen standardisierten oder auch individuellen Klassifikationssysteme können in eine MyCoRe-Anwendung integriert werden. Als Beispiel kann die Dewey-Dezimalklassifikation (DDC) genannt werden, die ein weit verbreitetes, international verwendetes Klassifikationssystem darstellt. Weitere Klassifikationen können für kategorisierende Angaben wie zum Beispiel Sprache, Schrift, Ressourcentypen oder Art des Materials erstellt werden. Anpassungen können in MyCoRe mit Hilfe des integrierten Klassifikationseditors vorgenommen werden. Die Bezeichnungen der Kategorien können in unterschiedlichen Sprachen vorliegen. Auch ist es möglich hierarchische Strukturen in einer Klassifikation abzubilden, so dass eine detaillierte Zuordnung vorgenommen werden kann. Eine Recherche im Bestand der digitalen Bibliothek kann dann auf zwei Wegen geschehen: zum einen als Suche nach Kategorien, zum anderen über die Navigation in einer Klassifikation.

MyCoRe bietet verschiedene Funktionen zur Suche in den Metadaten oder im Volltext von Dateien. Neben der gezielten Suche in bestimmten Datenfeldern, wie z.B. dem Titel von Publikationen, kann auch kombiniert in den Metadaten von Objekten gesucht werden, wie z.B. nach Stichwörtern im Datenfeld „Titel“ in Publikationen mit einem bestimmten Erscheinungsjahr. Darüber hinaus stellt MyCoRe eine eigene Suchanfragesprache zur Verfügung, die frei formulierbare Experten-Abfragen mit der *MyCoRe Query Language* (MCRQL) erlauben.

In digitalen Bibliotheken, die auf MyCoRe basieren, kann jeder erfasste Datensatz zu einem Objekt über *Persistent Identifier*¹¹ (URN) eindeutig referenziert werden. MyCoRe unterstützt die Generierung, Vergabe und Verwaltung von dauerhaften

¹¹ Siehe <http://www.persistent-identifier.de>

Identifikatoren im Namensraum „urn:nbn:de“, wie sie etwa die Deutsche Nationalbibliothek für Online-Publikationen einsetzt. Dokumente erhalten so eine dauerhafte, standardisierte und weltweit eindeutige Adresse, unter der sie erreichbar sind. Um die Integrität der gespeicherten Objekte zu sichern, kann in MyCoRe die zugehörige Prüfsumme berechnet und getestet werden. Vergebene URNs können entweder über eine OAI-Schnittstelle (siehe unten) mittels Epicur¹², oder über eine automatisch generierte E-Mail bei der Deutschen Nationalbibliothek registriert werden.

Schnittstellen

Ein großer Vorteil digitaler Bibliotheken ist die Möglichkeit des Datenaustauschs. Zu diesem Zweck bietet MyCoRe verschiedene Schnittstellen an, die technisch unterstützt werden. Da MyCoRe intern das XML-Format zur Speicherung von Metadaten verwendet, ist damit auch ein Datenaustausch einfach möglich. Metadaten können im XML-Format importiert und exportiert werden. Über *Java Servlets* sind diese XML-Daten auch externen Systemen via Hypertext-Übertragungsprotokoll (HTTP) zugänglich. Eine häufig benutzte Schnittstelle ist die OAI-Schnittstelle zum Einsammeln und Verteilen von Metadaten über das „*OAI Protocol for Metadata Harvesting*“ (OAI-PMH). Außerdem gibt es Zugänge, die es Suchmaschinen-Robots ermöglichen, die Metadaten und Volltexte der virtuellen Bibliothek zu indizieren und weltweit auffindbar zu machen. Auch das „Google Sitemap Protocol“ kann in MyCoRe-Anwendungen implementiert werden. Außerdem bietet MyCoRe eine SOAP-Schnittstelle für die Suche im Objektbestand, über die auch eine verteilte Suche über mehrere MyCoRe-Server konfiguriert werden kann, wie dies im Papyrus-Portal realisiert ist. Eine weitere bibliotheksspezifische Schnittstelle nennt sich z39.50¹³, über die einfache Suchanfragen im „z39.50 Prefix Query Format“ (PQF) gestellt und beantwortet werden. Über diese Schnittstelle kann z.B. ein auf MyCoRe basierender Dokumentenserver an einen Bibliothekskatalog angebunden werden.

Ausblick

Das breite Spektrum an Anwendungen zeigt, dass es möglich ist, virtuelle Bibliotheken für verschiedenste Bestände mit MyCoRe flexibel zu gestalten.

¹² Siehe <http://www.persistent-identifier.de/?link=220>

¹³ Siehe <http://www.loc.gov/z3950/agency>

Die Open-Source-Software MyCoRe eignet sich als Basis für Online-Bibliotheken und Repositorien mit Medien aller Art. Die Software wird kontinuierlich weiterentwickelt. Ein Ziel ist es, auch weiterhin die technischen Voraussetzungen zu schaffen, um die Kriterien des DINI Zertifikat für Dokumenten- und Publikationsservice (DINI 2010) erfüllen zu können.

Quellen der Abbildungen

Abb. 1: Bildschirmtastatur, Ausschnitt aus dem Onscreen-Keyboard in MyIHS, <http://myihs.dl.uni-leipzig.de/Keyboard.html>

Abb. 2: Bidirektionale Schrift, Ausschnitt aus der Refaiya-Bibliothek, http://www.refaiya.uni-leipzig.de/receive/RefaiyaBook_islamhs_00000038

Literatur

Arbeitsgruppe Elektronisches Publizieren DINI: DINI-Zertifikat Dokumenten- und Publikationsservice 2007, DINI-Schriften 3, April 2007 urn:nbn:de:kobv:11-10079197

Arbeitsgruppe Elektronisches Publizieren DINI: DINI-Zertifikat Dokumenten- und Publikationsservice 2010, DINI Schriften 3-de, Version 3.0, Juni 2010, urn:nbn:de:kobv:11-100109986

Freitag, S.; Gerhardt, M.; Kupferschmidt, J.; Scholl, R. : Papyrus Portal. Umsetzung einer Portallösung für alle Papyrussammlungen in Deutschland. Version 1.3, Leipzig, 04.09.2009, http://www.papyrusportal.de/PapyrusPortal_Dokumentation.pdf

Hanstein, T.; Karmun, A.; Kupferschmidt, J.: MyIHS. Basisprojekt zur datenbankgestützten Erschließung und digitalen Bereitstellung orientalischer Handschriften. Version 2.2.5., Leipzig, 06.08.2010
http://myihs.dl.uni-leipzig.de/Documentation_MyIHS.pdf

Hanstein, T.; Kupferschmidt, J.; Scharsky, A.: Digitalisierung fremdschriftlicher Texte bei Anwendung normierter Transkriptionssysteme. In: Ockenfeld, M. (Hrsg.): Generation international - die Zukunft von Information, Wissenschaft und Profession. Proceedings der 31. Online-Tagung der DGI. Frankfurt am Main 2009, S. 129-134.

Lützenkirchen, F. (2002): MyCoRe – Ein Open-Source-System zum Aufbau digitaler Bibliotheken. In: Datenbank Spektrum, 2(4), November 2002, 23-27, http://duepublico.uni-duisburg-essen.de/servlets/DerivateServlet/Derivate-11198/MyCoRe_DB-Spektrum.pdf

Oeltjen, W.: Metadaten-Management mit MyCoRe. In: Ball, R. (Hrsg.): Wissenschaftskommunikation der Zukunft, Schriften des Forschungszentrums Jülich, Band 18, 2007, S. 185-191

Alle Web-Adressen wurden am 6.9.2010 geprüft.